Interpretation of Single and Serial Measures of HE4 and CA125 in Asymptomatic Women at High Risk for Ovarian Cancer

Nicole Urban1, Jason Thorpe1, Beth Y. Karlan3, Martin W. McIntosh1, Melanie R. Palomares4, Mary B. Daly5, Pamela Paley2, and Charles W. Drescher1,2

Abstract

Background: Human epididymis protein 4 (HE4) is approved for clinical use with CA125 to predict epithelial ovarian cancer in women with a pelvic mass or in remission after chemotherapy. Previously reported reference ranges for HE4 are inconsistent.

Methods: We report positivity thresholds yielding 90%, 95%, 98%, and 99% specificity for age-defined populations of healthy women for HE4, CA125, and Risk of Ovarian Malignancy Algorithm (ROMA), a weighted average of HE4 and CA125. HE4 and CA125 were measured in 1,780 samples from 778 healthy women aged >25 years with a documented deleterious mutation, or aged >35 years with a significant family history. Effects on marker levels of a woman’s age, ethnicity, and epidemiologic characteristics were estimated, as were the population-specific means, variances, and within- and between-woman variances used to generate longitudinal screening algorithms for these markers.

Results: CA125 levels were lower with Black ethnicity (P = 0.008). Smoking was associated with higher HE4 (P = 0.007) and ROMA (P < 0.019). Continuous oral contraceptive use decreased levels of CA125 (P = 0.041) and ROMA (P = 0.12). CA125 was lower in women age >55, and HE4 increased with age (P < 0.01), particularly among women age >55.

Conclusions: Because of the strong effect of age on HE4, thresholds for HE4 are best defined for women of specific ages. Age-specific population thresholds for HE4 for 95% specificity ranged from 41.4 pmol/L for women age 30 to 82.1 pmol/L for women age 80.

Impact: Incorporation of serial marker values from screening history reduces personalized thresholds for CA125 and HE4 but is inappropriate for ROMA. Cancer Epidemiol Biomarkers Prev; 21(11); 2087–94. ©2012 AACR.

Introduction

Human epididymis protein 4 (HE4) is an epithelial ovarian cancer (EOC) serum marker that was identified originally as a candidate early detection marker (1). Since then it has been shown to be potentially useful for remission monitoring (2–4) and has been cleared by the U.S. Food and Drug Administration (FDA) for that use. It has also been shown to perform as well or better than CA125 in distinguishing benign from malignant tumors in women with a pelvic mass (5–7). The Risk of Ovarian Malignancy Algorithm (ROMA) which includes CA125 as well as HE4 (8) has been cleared by the FDA for evaluation of a pelvic mass. HE4 has also been reported to contribute to identification of malignancy in women with nonspecific abdominal/pelvic symptoms (9). Most recently, it has been reported to provide detectable signal about a year before diagnosis (10, 11), renewing interest in it as an early detection marker. We have recently reported that HE4 outperforms imaging as a second-line screen when rising CA125 is used as a first-line screen (12).

The efficacy of screening for EOC has not been shown. Use of both imaging and CA125 (using a threshold of 35 U/mL for all women) annually in women aged 55 to 74 leads to unnecessary surgery (13) without reducing mortality (14). However, a multimodal strategy using rising CA125 annually to select women for imaging is yielding acceptable positive predictive value (PPV) in an efficacy trial in the United Kingdom where the longitudinal Risk of Ovarian Cancer Algorithm (ROCA) (15) is used to measure rising CA125 (16). The ROCA has not been applied to HE4. An alternative strategy for using novel
markers in a longitudinal algorithm is the Parametric
Empirical Bayes (PEB) longitudinal algorithm (17).
Use of the PEB algorithm has been reported for CA125 previ-
ously (17, 18), and it can be readily adapted for use with
HE4 because it is fit using data from healthy women only.

HE4 is in clinical use but reference ranges for HE4 have
been reported only very recently and remain uncertain
because of inconsistency. Park and colleagues report the
97.5% upper reference limit for HE4 to be 33.2 pmol/L for
a young Asian population (19), compared with the
median value of 48 pmol/L in healthy controls reported
by the same authors for an Asian hospital population (20).

Moore and colleagues report the upper 95th percentile
of HE4 to be 89 pmol/L for premenopausal women, 128
pmol/L for postmenopausal women, and 115 pmol/L for
all women, lower in pregnancy and rising with age (21).

Other investigators have used positivity thresholds for
HE4 of 70 pmol/L (22) and 150 pmol/L (23) in clinical
validation studies.

Here we report HE4 positivity thresholds yielding 90%,
95%, 98%, and 99% specificity, that account for age,
derived from analysis of 1,780 samples from 778 healthy
women aged ≥25 years with a documented deleterious
mutation, or aged ≥35 years with a significant family
history. We also report parameters characterizing the
within- and between-components of variance for HE4 to
allow researchers to calculate the thresholds for HE4
when adjusting for a woman’s screening history as well
as her age using the PEB rule (18). Whenever the total
variance of a marker is dominated by between-women
differences, the PEB decision rule will maintain overall
test specificity but may achieve better sensitivity, longer
lead time, and lower thresholds for the majority of women
compared with a single threshold (ST) rule (18, 24).

Interpretation of serial measures of HE4 using a longitudinal
algorithm has not been previously reported. For compar-
ison we also report the PEB-relevant parameters for
CA125 and ROMA in this population.

Materials and Methods

Women enrolled in the Novel Markers Trial (NMT)
through February 29, 2012 were eligible for the current
study if at the time of enrollment they were aged ≥25 years
and reported having tested positive for a deleterious
mutation in BRCA1 or BRCA2, or aged ≥35 years with
a significant family history and were not diagnosed with
EOC during follow-up. The NMT is a 2-arm randomized
multi-institutional Phase I screening trial sponsored by
the National Cancer Institute (25). The NMT introduces
HE4 as either a first- or second-line screen in a multimodal
screening strategy that includes CA125, HE4, and trans-
vaginal ultrasound (TVU) (26). Participants completed a
baseline questionnaire that provided information about
ovarian cancer family history and other risk factors, and
contributed up to 5 blood samples about 6 months apart.

HE4 and CA125 were measured in 1,780 samples from 778
healthy women on the Abbott Architect automated plat-
form in a Clinical Laboratory Improvement Amend-
ments-approved laboratory using FDA-approved kits.
The ROMA Predictive Index (PI) was calculated using
previously defined pre- and post-menopause formulae
(8). We used $PI = -12.0 + 2.38 \cdot \ln(\text{HE4}) + 0.626 \cdot \ln$(CA125)$ for women under age 50, and$PI = -8.09 + 1.94 \cdot \ln(\text{HE4}) + 0.732 \cdot \ln$(CA125)$ for the remaining women.

To address interpretation of single and longitudinal
measures of markers, we first stratified ages of all women
by 4 clinically relevant age strata: age <45, 45–65, age <55,
55< age <65, and age ≥65. Natural-log (ln) transformed
marker values were used in all analyses and then trans-
formed back to their raw scale for reporting purposes.

Because the ROMA PI is a linear combination of log-
transformed CA125 and HE4 values (8), no additional
transformations were needed and results are reported on
its raw scale.

A linear regression model associating marker concen-
tration ($Y$) to covariates was fit using a generalized esti-
mating equations clustering (27) on participant identifier
to adjust for varying numbers of observations per partic-
ent to estimate the model $Y = \beta_0 + \beta_1X_1 + \ldots + \beta_nX_n + \beta + \epsilon$, where $\epsilon$ is a residual normal distribution
with mean 0 and residual variance $V$. The coefficients
were used to determine the magnitude of effect for each
epidemiologic covariate. The regression predictors
included epidemiologic covariates, and main effects
of the age strata, age as a continuous variable, and also an
interaction between age and age stratum to allow the
effect of age to differ among the different age-defined
strata. The regression line was used to produce reference
ranges for HE4, which has a strong trend with age within
each age stratum. Age-defined population thresholds are
calculated from the fitted values and residual standard
error on the basis of an assumption that the residual errors
are approximately normally distributed. For CA125,
where levels are constant in the younger and older wom-
ên- age-defined population thresholds were determined
from empirically defined percentiles of the marker within
those age ranges.

The marker threshold for an individual woman can be
personalized if screening history is known, including age
and marker value at each screen. To control for screening
history we used the PEB screening rule, taking advantage
of its ability to accommodate covariates (17, 18). We
assume that each woman’s marker levels may systemat-
ically differ, on average by a constant amount, from the
expected value of $Y$, where $Z$ is the residual normal distribution
with mean 0 and residual variance $V$. The coefficients
were used to determine the magnitude of effect for each
epidemiologic covariate. The regression predictors
included epidemiologic covariates, and main effects
of the age strata, age as a continuous variable, and also an
interaction between age and age stratum to allow the
effect of age to differ among the different age-defined
strata. The regression line was used to produce reference
ranges for HE4, which has a strong trend with age within
each age stratum. Age-defined population thresholds are
calculated from the fitted values and residual standard
error on the basis of an assumption that the residual errors
are approximately normally distributed. For CA125,
where levels are constant in the younger and older wom-
en- age-defined population thresholds were determined
from empirically defined percentiles of the marker within
those age ranges.

The marker threshold for an individual woman can be
personalized if screening history is known, including age
and marker value at each screen. To control for screening
history we used the PEB screening rule, taking advantage
of its ability to accommodate covariates (17, 18). We
assume that each woman’s marker levels may systemat-
ically differ, on average by a constant amount, from the
expected value of $Y$, where $Z$ is the residual normal distribution
with mean 0 and residual variance $V$. The coefficients
were used to determine the magnitude of effect for each
epidemiologic covariate. The regression predictors
included epidemiologic covariates, and main effects
of the age strata, age as a continuous variable, and also an
interaction between age and age stratum to allow the
effect of age to differ among the different age-defined
strata. The regression line was used to produce reference
ranges for HE4, which has a strong trend with age within
each age stratum. Age-defined population thresholds are
calculated from the fitted values and residual standard
error on the basis of an assumption that the residual errors
are approximately normally distributed. For CA125,
where levels are constant in the younger and older wom-
en- age-defined population thresholds were determined
from empirically defined percentiles of the marker within
those age ranges.

The marker threshold for an individual woman can be
personalized if screening history is known, including age
and marker value at each screen. To control for screening
history we used the PEB screening rule, taking advantage
of its ability to accommodate covariates (17, 18). We
the regression equation and the within- and between-woman variance parameters using the following formula

\[
\text{Threshold} = 
\left( \frac{Z(B_0)}{\sqrt{1 - BB_n}} \right) + b_0 + b_1 X_1 + b_2 X_2 + \cdots
\]

where \( C_n \) is the alpha quantile, a standardized normal distribution (e.g., 1.64 for 95% specificity). See McIntosh and Urban (17) for derivation and details. Parameters required for calculating the PEB thresholds for HE4 and CA125 are provided in the Results section. With longer screening histories, \( B_n \) increases to a maximum of 1 and the intercept term equals a woman’s unique individual value level \( Z \) When “\( n = 0 \)” then \( B_0 = 0 \) and the PEB threshold defaults to the population-defined linear regression values. The person-specific residual variance also decreases as screening history accumulates, meaning that the reference range for a woman shrinks in size over time. Theoretical results show that at comparable population-wide specificities, the PEB uniformly improves upon rules that ignore screening history and improves over simpler rules that make decisions on the basis of simple change-from-baseline (17, 18).

**Results**

Baseline characteristics of NMT participants included in this report are reported in Table 1. The majority of women were aged 45 to 64, Caucasian, non-Hispanic, parous, non-smoking, non-hysterectomized and non-users of hormone replacement therapy (HRT), and most did not have a prior tubal ligation. A deleterious mutation in a \( BRCA1 \) or \( BRCA2 \) gene was reported by 17.7% of women included in this study. A personal history of breast cancer was reported by 17.2% and a family history of breast cancer was reported by 75.3% of women. A family history of ovarian cancer was reported by 43.1%, and Ashkenazi Jewish lineage was reported by 19.5% of women. Current oral contraceptive (OC) use was reported by 5.0% of women; half of these reported continuous OC use resulting in cessation of periods.

Effects of age and age-adjusted effects of other covariates on mean marker levels are reported in Table 2. We first evaluated the effect of age on the marker levels within each age stratum. CA125 levels are lower for women with age \( \geq 55 \) than in age \( <45 \) (\( P < 0.001 \)), but levels are constant and unchanging with respect to age within these categories (\( P = 0.49 \) for age \( <45 \) and \( P = 0.797 \) for age \( \geq 55 \)). Rapid changes in CA125 were found in the strata defined by \( 45 \leq \text{age} < 55 \), with a decline of 30% over the 10-year period (\( P = 0.006 \)).

For HE4 an increasing trend was found in all age strata but the most dramatic change was found in women after age 55. Before age 45, Ln HE4 concentrations elevate by an average of 0.0082 per year, or exp \( \left( 0.0082 \times 10 \right) = 10.04 \) per decade on the raw scale (\( P = 0.013 \), but they elevate by 0.0158 per year faster than the reference group, or exp \( \left( 0.0158 + 0.0082 \right) \times 10 \) = 10.27% per decade on the raw scale after age 55. Although the slope among women with \( 45 \leq \text{age} < 55 \) was just 2.7%, it did not differ significantly from that of women age \( <45 \) (\( P = 0.40 \)), suggesting that these age categories could potentially be combined when interpreting HE4. An analysis combining these 2 younger periods finds HE4 with a slope of 0.00602, or 6.2% change per decade on the raw scale before age 55. We did not combine these groups in our analyses, however.

For ROMA, which includes both CA125 and HE4, an increasing trend was apparent in the reference group (age \( <45 \), \( P = 0.012 \)), and the slope was higher among the older age groups though the difference was only significant in women \( 45 \leq \text{age} < 55 \) (\( P < 0.005 \)). However, the definition of ROMA, especially because its definition is different for women before and after menopause (age 50 here), is difficult to interpret the reference ranges of ROMA over time. We provide it here for reference.

### Table 1. Number and percentage of women and samples included in analyses by age, ethnicity, and personal characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Participants</th>
<th>Samples contributed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total count</td>
<td>778 (100.0%)</td>
<td>1780 (100.0%)</td>
</tr>
<tr>
<td>Age (mean, SD)</td>
<td>52.3 (10.9)</td>
<td>53.5 (10.9)</td>
</tr>
<tr>
<td>Age (&lt;45)</td>
<td>190 (24.4%)</td>
<td>370 (20.8%)</td>
</tr>
<tr>
<td>(45 \leq \text{Age} &lt; 55)</td>
<td>276 (35.5%)</td>
<td>635 (35.7%)</td>
</tr>
<tr>
<td>(55 \leq \text{Age} &lt; 65)</td>
<td>215 (27.6%)</td>
<td>513 (28.8%)</td>
</tr>
<tr>
<td>Age (\geq65)</td>
<td>97 (12.5%)</td>
<td>262 (14.7%)</td>
</tr>
<tr>
<td>Non-White race</td>
<td>78 (10.0%)</td>
<td>159 (8.9%)</td>
</tr>
<tr>
<td>Black</td>
<td>10 (1.3%)</td>
<td>18 (1.0%)</td>
</tr>
<tr>
<td>Asian</td>
<td>22 (2.8%)</td>
<td>42 (2.4%)</td>
</tr>
<tr>
<td>Other</td>
<td>46 (5.9%)</td>
<td>99 (5.6%)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>36 (4.6%)</td>
<td>66 (3.7%)</td>
</tr>
<tr>
<td>Ashkenazi Jewish</td>
<td>152 (19.5%)</td>
<td>325 (18.3%)</td>
</tr>
<tr>
<td>Parous</td>
<td>520 (66.8%)</td>
<td>1,186 (66.6%)</td>
</tr>
<tr>
<td>Hysterectomy</td>
<td>67 (8.6%)</td>
<td>156 (8.8%)</td>
</tr>
<tr>
<td>Tubal ligation</td>
<td>103 (13.2%)</td>
<td>253 (14.2%)</td>
</tr>
<tr>
<td>Current smoker</td>
<td>26 (3.3%)</td>
<td>58 (3.3%)</td>
</tr>
<tr>
<td>Current HRT use</td>
<td>101 (13.0%)</td>
<td>252 (14.2%)</td>
</tr>
<tr>
<td>Current OC use</td>
<td>39 (5.0%)</td>
<td>91 (5.1%)</td>
</tr>
<tr>
<td>Discontinuous OC</td>
<td>21 (2.7%)</td>
<td>48 (2.7%)</td>
</tr>
<tr>
<td>Continuous OC</td>
<td>18 (2.3%)</td>
<td>43 (2.4%)</td>
</tr>
<tr>
<td>Family history of breast cancer</td>
<td>586 (75.3%)</td>
<td>1340 (75.3%)</td>
</tr>
<tr>
<td>Family history of ovarian cancer</td>
<td>335 (43.1%)</td>
<td>792 (44.5%)</td>
</tr>
<tr>
<td>Personal history of breast cancer</td>
<td>134 (17.2%)</td>
<td>254 (14.3%)</td>
</tr>
<tr>
<td>Deleterious mutation</td>
<td>138 (17.7%)</td>
<td>261 (14.7%)</td>
</tr>
</tbody>
</table>
Among all other covariates, concentrations of CA125 were lower in women with Black ethnicity (26.6%; \( P = 0.008 \)) and with continuous OC use (18.9%; \( P = 0.041 \)). Concentrations of HE4 were higher in current smokers (21%; \( P = 0.007 \)). ROMA was higher in smokers (\( P < 0.019 \)) and lower in women reporting continuous OC use (\( P = 0.012 \)).

Table 3 reports age-defined population thresholds. For CA125, where concentrations are constant within the age ranges \( < 45 \) and \( \geq 55 \), a single reference range is appropriate for all women in these age ranges, and thresholds are determined empirically and without distributional assumptions. However, for HE4, because it changes so dramatically with age we provide the thresholds at specific ages predicted from the regression equation. However, because the thresholds are linear on the log scale, one can predict reference ranges for any intermediate age by using linear interpolation on the appropriate scale; for example, the 95th percentile for age 65 is \( \exp \left[ 0.5 \ln (50.8) + 0.5 \ln (64.6) \right] = 57.3 \). Note that only age is accounted for in Table 3. Applying the appropriate percentage reduction using the effect sizes from Table 2 can approximate adjustments for other covariates (e.g., a 21% increase in HE4 for a current smoker). However, because sample sizes for the significant predictors of CA125 and HE4 are small, one must

### Table 2. Effect of age, ethnicity, and personal characteristics on marker concentrations

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Covariate</th>
<th>Sample size</th>
<th>CA125</th>
<th>HE4</th>
<th>ROMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age-related effects on ln marker concentrations(^a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td>370</td>
<td>2.39859</td>
<td>&lt;0.001</td>
<td>2.99346</td>
</tr>
<tr>
<td>Age (&lt; 45) (Reference)</td>
<td></td>
<td>635</td>
<td>1.92218</td>
<td>0.007</td>
<td>0.25002</td>
</tr>
<tr>
<td>Age (\geq 55)</td>
<td></td>
<td>1.015</td>
<td>-0.35088</td>
<td>0.365</td>
<td>-0.98948</td>
</tr>
<tr>
<td>Trends</td>
<td></td>
<td>370</td>
<td>0.00524</td>
<td>0.490</td>
<td>0.00820</td>
</tr>
<tr>
<td>Age (&lt; 45) (Reference)</td>
<td></td>
<td>635</td>
<td>-0.04155</td>
<td>0.006</td>
<td>-0.00551</td>
</tr>
<tr>
<td>Age (\geq 55)</td>
<td></td>
<td>775</td>
<td>-0.00224</td>
<td>0.797</td>
<td>0.01579</td>
</tr>
<tr>
<td>Covariate effects on percentage change of marker concentrations</td>
<td></td>
<td></td>
<td></td>
<td>% Change</td>
<td>( P ) value</td>
</tr>
<tr>
<td>Age-adjusted effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parous</td>
<td></td>
<td>1,186</td>
<td>4.0%</td>
<td>0.355</td>
<td>0.4%</td>
</tr>
<tr>
<td>Non-White race</td>
<td></td>
<td>159</td>
<td>6.3%</td>
<td>0.371</td>
<td>3.6%</td>
</tr>
<tr>
<td>Black</td>
<td></td>
<td>18</td>
<td>-26.6%</td>
<td>0.008</td>
<td>-3.3%</td>
</tr>
<tr>
<td>Asian</td>
<td></td>
<td>42</td>
<td>20.4%</td>
<td>0.112</td>
<td>7.7%</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>99</td>
<td>7.5%</td>
<td>0.417</td>
<td>2.9%</td>
</tr>
<tr>
<td>Hispanic</td>
<td></td>
<td>66</td>
<td>19.10%</td>
<td>0.097</td>
<td>1.8%</td>
</tr>
<tr>
<td>Ashkenazi Jewish</td>
<td></td>
<td>325</td>
<td>-7.4%</td>
<td>0.163</td>
<td>0.9%</td>
</tr>
<tr>
<td>Prior Tubal ligation</td>
<td></td>
<td>253</td>
<td>0.4%</td>
<td>0.935</td>
<td>-0.3%</td>
</tr>
<tr>
<td>Current smoker</td>
<td></td>
<td>58</td>
<td>0.3%</td>
<td>0.977</td>
<td>21.0%</td>
</tr>
<tr>
<td>Current OC use</td>
<td></td>
<td>91</td>
<td>-13.0%</td>
<td>0.100</td>
<td>-1.8%</td>
</tr>
<tr>
<td>Discontinuous OC</td>
<td></td>
<td>48</td>
<td>-5.4%</td>
<td>0.641</td>
<td>2.4%</td>
</tr>
<tr>
<td>Continuous OC</td>
<td></td>
<td>43</td>
<td>-18.9%</td>
<td>0.041</td>
<td>-5.8%</td>
</tr>
<tr>
<td>Personal history of breast cancer</td>
<td></td>
<td>584</td>
<td>-3.0%</td>
<td>0.497</td>
<td>3.9%</td>
</tr>
</tbody>
</table>

\(^a\)Coefficients and \( P \) values obtained from a linear regression using the formula: \( \text{ln(marker level)} = \text{intercept} + \text{age} + [I(45 < \text{age} < 55) + \text{age}] + [I(45 < \text{age} < 55) + \text{age} + I(\text{age} > 55)] \), where indicator variables \( I(\ldots) \) are set to 1 when the condition is met and 0 otherwise. Together, the intercept \( [e.g., I(45 < \text{age} < 55)] \) and trend \( [e.g., \text{"age" \[I(45 < \text{age} < 55)\]} \) terms allow for independent linear trends to be estimated within each age range.

NOTE: The reference group is women under age 45. Age-related effects on ln(marker levels) are presented in the top half of the table. The bottom half presents the effects of other covariates in terms of percentage change of marker levels. The marker value expected for an individual woman can be calculated using this table using instructions in the table footnote. Sample size refers to blood draws not number of women. To approximate the marker value expected for a woman of a specific age, first approximate the baseline LN marker level for the woman’s age as follows. Add the intercept terms (reference group plus the term from the age appropriate category) to the age multiplied by the age-related trend (reference group trend + age appropriate category trend). For example, baseline LN CA 125 level for a woman of age 50 will equal \( Y = (2.39859 + 1.92218) + 50 \times 0.00524 = 2.50 \), or 12.23 U/mL on the raw scale. To adjust the level on the raw scale for covariates, increase or decrease the calculated raw scale using the approximated effect size for that covariate. For example, a woman of age 50 who is of Black ethnicity will be expected to have CA 125 levels 26.6% lower, or 3.25 U/mL on the raw scale.
be cautious as one may anticipate a wide confidence interval around their effects.

For HE4, age-specific population thresholds for 95% specificity ranged from 41.4 pmol/L for women age 30 to 82.1 pmol/L for women age 80. The magnitude of the effect of age on HE4 means and thresholds can be seen in Table 3. The 10-year span from age 30 to 40 increases the 95th percentile for HE4 by only 4 points, but a 10-year span from age 60 to 70 increases HE4 by nearly 14 points. The effect of age on HE4 means and thresholds can be seen in Table 3. The 10-year span from age 30 to 40 increases the 95th percentile for HE4 by only 4 points, but a 10-year span from age 60 to 70 increases HE4 by nearly 14 points. The association with age for HE4 has been previously reported, but not the distinct behavior on either side of 55 years (12, 21, 29). ROMA thresholds similarly increase with age. For CA125, thresholds yielding 95% specificity are reported, but not the distinct behavior on either side of 55 years (12, 21, 29). ROMA thresholds similarly increase with age. For CA125, thresholds yielding 95% specificity are reported, but not the distinct behavior on either side of 55 years (12, 21, 29).

Incorporation of serial marker values from screening history potentially reduces thresholds for CA125 and HE4 but was found to be inappropriate for ROMA. Thresholds personalized for both covariates and screening history are obtained using the PEB algorithm for Ln(HE4) and Ln(CA125). We do not include ROMA in the PEB calculation, as it is not recommended for use with healthy women for which longitudinal screening algorithms will be applied, and because applying a longitudinal algorithm to the ROMA as a means to combine markers is not necessarily optimal or superior to applying the PEB algorithms to CA125 and HE4 separately and defining positivity as either marker positive. The PEB parameters $V$ and $B$, from the equation, are shown in Table 4 [the ICC (or $B$) is equal to $B_n$ when $n = 1$. We also provide $B_n$ for scenarios when screening history is between 0 and 6 screens. These values and the results of Table 2 can be used to calculate the threshold for any screening history for any woman. For HE4, the residual variance $V$ and $B$ is computed from the regression equation represented in the footnote to Table 2. However, for CA125 these values were computed separately using data within each age stratum, and parameters are reported separately for each stratum. Note because CA125 values change so dramatically in the 45 $< age < 55$ stratum, it is not recommended to use the PEB rule during this time, but we report them here for completeness. The ICC for HE4 is smaller than for CA125 suggesting that its screening history is less informative than that of CA125, but its value exceeds 0.5, suggesting that personalizing thresholds may have a meaningful effect compared with ignoring screening history (18, 24).

The effectiveness of controlling for screening history is shown by the "% reduction in offset" in Table 4, referring to the percentage that the "offset" in the PEB expression, which represents the deviation from the expected value that is tolerated before a screen is declared positive, was reduced. The goal of the PEB rule is to provide a narrower predicted reference for each woman while maintaining overall test specificity. The narrower reference range will facilitate earlier detection of a tumor: a reference range that is half the width can detect tumors when markers elevate half the amount. As seen in Table 4, by the 4th screen the reference range for HE4 is on average 30% narrower when controlling for screening history than when ignoring prior marker values. CA125, because of its higher ICC, can achieve an even greater reduction in reference range size. Note that the ICC compares screening rules that ignore history to those using the same marker accounting for screening history; the ICC does
not measure the quality of the marker otherwise. As seen in Table 4, the incremental benefit of controlling for screening history diminishes over time. For CA125 with women age ≥55, 2 previous screens can reduce the offset by 43% but a 49% reduction is not obtained until 6 screens.

Discussion

Interpretation of single and serial measures of HE4 and CA125 in women at high risk for ovarian cancer can be summarized in thresholds for positivity at first and subsequent screens. We investigated characteristics that might affect population thresholds for use at first screens. For HE4, we found concentrations to depend on age, with a quite dramatic change over time after age 55. The thresholds reported by us and others [e.g., Moore and colleagues, (21) or Park and colleagues (20)] are difficult to compare owing to differences in the ages of the populations, in addition to potential differences in race. We have shown that adequately controlling for effects of age on HE4, given its rapid rise after age 55, will require reference ranges that depend on an individual woman’s specific age. This is different from CA125 where all women in broad age categories share the same CA125 mean levels. For CA125, if 95% specificity is desired population thresholds are 31.8 U/mL and 22.2 U/mL for women age <45 and ≥55, respectively. For HE4, it is useful to categorize women more finely because HE4 increases steadily with age. The age-specific population thresholds for HE4 for 95% specificity nearly double over the age range studied, increasing from 41.4 pmol/L for a 30-year-old woman to 82.1 pmol/L for an 80-year-old woman. Although our analysis focused on high-risk women, it is likely to be generalizable to low-risk women because most of the high-risk women in this study were at modestly increased risk, and because prior work suggests that risk does not influence marker levels (30).

Reference ranges for ROMA may provide a means to screen women by combining CA125 and HE4. However, although ROMA PI was derived using statistically optimal procedures for use in a cross-sectional study (8) that are appropriate for use there, its optimality for use in a longitudinal study has not been shown. It may be inferior to other alternative approaches such as using the markers separately then making decisions on the basis of whether one or both are elevated. In its current form, ROMA is not appropriate for use in early detection among asymptomatic women.

A limitation of this study is that the healthy women included in the NMT were not routinely characterized for the presence or absence of uterine leiomyoma, which may cause elevation in CA125. To address this possibility, we reviewed medical charts as well as imaging reports for all 53 women who had NMT-protocol-indicated TVU to identify any leiomyoma occurring in study participants with elevated markers. Four such women were identified. In all 4, TVU was performed because of transitory elevation in CA125. The presence of a leiomyoma appears to increase within-woman variability in CA125 and to cause occasional elevation above 99% thresholds.

Results of this study are more relevant to research than to clinical application of markers because screening for ovarian cancer is currently not recommended. It is especially important that primary care physicians understand that there is currently no evidence to support screening, especially using imaging which is often their preferred modality (31). The Prostate Lung Colon and Ovary (PLCO) trial in the United States failed to demonstrate a reduction in EOC-specific mortality with annual screening using both CA125 and TVU concurrently (14). At the prevalence screen in the PLCO trial, rates of surgery were 2.6 times higher for TVU-positive only than for CA125-positive only (32). Reports from the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKC-TOCS) suggest that CA125 may outperform TVU in screening for EOC when a longitudinal algorithm is used to measure rising CA125 (16). The sensitivity, specificity,
and PPV for all primary invasive EOCs identified at the prevalence screen in the UKCTOCS were 89.5%, 99.8%, and 35.1%, respectively, for the multimodal strategy using CA125 to select women for TVU, and 75.0%, 98.2%, and 2.8%, respectively, for TVU alone (16). Although results for the CA125 arm are promising, it is not yet known whether the sequential multimodal screening strategy will reduce mortality. The clinical benefit of using HE4 and the PEB algorithm for ovarian cancer screening has also not been documented. Data on clinical outcomes of the NMT cohort are currently being gathered. HE4 and CA125 interpreted by a longitudinal algorithm may have a role to play in EOC screening trials in the future. If research were to identify markers with improved operating characteristics, the population of patients with high risk (mutation carriers and those with significant pedigrees) might one day benefit.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors’ Contributions

Conception and design: N. Urban, J.D. Thorpe, B.Y. Karlan, M.R. Palomares, C.W. Drescher

Development of methodology: N. Urban, J.D. Thorpe, B.Y. Karlan, M.W. McIntosh

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): N. Urban, B.Y. Karlan, M.R. Palomares, M.B. Daly, P.J. Foley, C.W. Drescher

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): J.D. Thorpe, B.Y. Karlan, M.W. McIntosh, M.R. Palomares, M.B. Daly, C.W. Drescher

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): J.D. Thorpe, B.Y. Karlan

Writing, review, and/or revision of the manuscript: N. Urban, J.D. Thorpe, B.Y. Karlan, M.W. McIntosh, M.R. Palomares, M.B. Daly, P.J. Foley, C.W. Drescher

Study supervision: N. Urban, B.Y. Karlan, M.R. Palomares

Acknowledgments

The authors gratefully acknowledge helpful comments provided during review of manuscript from Beth Schodin and Barry Dowell of Abbott Diagnostics. The authors also gratefully acknowledged the support for the Translational and Outcomes Research Laboratory from NIH/NCI U01 CA152637 and the Canary Foundation, support for clinical centers from the Marsha Rivkin Center for Ovarian Cancer Research and Canary Foundation, and a grant of no-charge study materials from Abbott Laboratories. The content is solely the responsibility of the authors and does not necessarily represent official views of NIH/NCI, Canary Foundation, Marsha Rivkin Center for Ovarian Cancer Research or Abbott Laboratories.

Grant Support

This work was supported by the Pacific Ovarian Cancer Research Consortium, Award Number P50 CA083636 from the NIH/National Cancer Institute (NCI). NIH/NCI P50 CA083636 and U01 CA152637.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 18, 2012; revised July 24, 2012; accepted August 27, 2012; published OnlineFirst September 7, 2012.

References

Interpretation of Single and Serial Measures of HE4 and CA125 in Asymptomatic Women at High Risk for Ovarian Cancer

Nicole Urban, Jason Thorpe, Beth Y. Karlan, et al.


Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-12-0616

This article cites 30 articles, 9 of which you can access for free at:
http://cebp.aacrjournals.org/content/21/11/2087.full.html#ref-list-1

This article has been cited by 3 HighWire-hosted articles. Access the articles at:
/content/21/11/2087.full.html#related-urls

Sign up to receive free email-alerts related to this article or journal.

To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.